

The background of the entire page is a light teal network diagram consisting of interconnected nodes and lines. The word "RESOLVE" is written in a large, bold, dark teal font, with "NETWORK" in a smaller, all-caps, dark teal font directly below it. A thin black diagonal line is positioned to the right of the word "RESOLVE".

RESOLVE

NETWORK

Going Back to Basics: Research on Violent Extremism

Kris Inman

The views in this publication are those of the author alone. They do not reflect the views of the RESOLVE Network, its partners, the U.S. Institute of Peace, or the United States Government or any of its constituent parts.

ABSTRACT

This chapter reviews research fundamentals, from formulating a research question to research ethics, which are easy to overlook in applied research conducted to understand countering violent extremism (CVE). The chapter offers examples of good and bad research practices, as well as practical ways to improve research designs by going back to the basic principles for both qualitative and quantitative data collection and analysis. While the chapter can serve as a practical guide for sound research design, it goes beyond the nuts-and-bolts of research and discusses why these research best practices are important for policymakers and for moving our knowledge of CVE forward.

INTRODUCTION

It is exciting and rewarding, but also challenging, to work in the field of countering violent extremism (CVE). As a relatively new field at only about 10 years old, we face a lot of pressure to implement effective CVE responses as violent extremism (VE) around the world continues to increase, despite the kinetic counter-terrorism efforts of the past 20 years. We are continually in search of evidence about what works and what does not, what is scalable and sustainable, and what unintended consequences lurk behind best intentions. In the face of rapidly evolving VE groups like the Islamic State, Boko Haram, Al Qaeda, Al Shabaab, and other local terrorist activities around the world, it can seem like we are always at least one step behind and constantly re-creating the wheel in our responses to VE. One area where we can avoid the latter, and thereby address the former, is in how we conduct research on CVE. Research best practices and principles are well-established and transfer across sectors and disciplines. This chapter reviews some of those best practices that, when adopted in research, improve the capacity of applied researchers and analysts to support CVE policy and programs and other peacebuilding efforts.

Studying and responding to VE requires innovative and flexible approaches that can deal with the complexity of VE dynamics. However, when innovating in the research sphere, applied researchers and analysts often forget the basic tenants of research design. These tenants include formulating a non-normative question that answers “who,” “how,” “why,” “where,” or “what”; clearly defining concepts in the research question; designing a research approach that appropriately measures those concepts; systematically collecting and analyzing data to answer the research question with evidence; clearly discussing inference based on the evidence; and respecting ethical standards.

The power of going back to basics, even for experienced researchers, should not be underestimated. While innovation is necessary in research on VE, innovation must build on a solid foundation of basic research techniques. Failure to hold on to the basics inhibits practitioners from building the body of knowledge required to develop a sophisticated understanding of VE, and thereby hampers advancing responses to VE. Moreover, it hinders us from providing solid evidence to policy makers.

FORMULATING THE RESEARCH QUESTION

The strength of the body of knowledge around a topic like VE depends on the strength of the individual research projects that contribute to that knowledge. The strength of an individual research project begins with the formulation of a good research question. Good research questions are clearly articulated; non-normative, asking about what **is**, rather than about what **should be**; and clearly establish whether the research in question is intended to describe, explain, predict, or proscribe the topic. How many CVE reports and research articles can you recall that clearly articulate a research question? Chances are if you take time to reflect on this question, you will find that many reports and research articles you can think of do not ask a research question.

Pick an article on VE or CVE and see if you can identify a question mark in the introduction. If there is a clearly articulated research question, is it a good one? Is it researchable? In our field, a lot of research questions are formulated as “yes/no” questions, such as “Do women support Boko Haram?” or “Is poverty a driver of violent extremism?” Yes or no questions do not meet the standards of good research questions because research questions need to orient the audience to whether the research is about describing, explaining, predicting, or proscribing the phenomena in question.

One way to reformulate the question “Do women support Boko Haram,” is to reflect on what you as the researcher (or as a policymaker, donor, or implementer) want to know about the relationship between women and Boko Haram. Do you want to know **whether** some women support Boko Haram? If so, a better way to formulate the question is “In what ways do women in Northern Nigeria support Boko Haram?” and possibly also “what are the demographic characteristics of women who support Boko Haram?” Or, do you want to know **why** some women support Boko Haram? In this case, a better way to ask the question is simply “why do some women in Northern Nigeria support Boko Haram?” Do you want to **compare** women’s support to Boko Haram across the territories within which Boko Haram is active? If so, then you may ask a research question like “How does women’s support of Boko Haram in the Lake Chad Basin differ for support in Northern Nigeria and why?”

Related to posing a good research question is posing an appropriate number of questions. Good, solid research asks one question (or at most two closely related questions) at a time and then thoroughly investigates that question. One of the first things I learned as a graduate student in methods classes was the KISS principle: Keep it Simple, Stupid. This applied to several topics in methods, but the very first one was the number of research questions we asked in our research. One research project cannot do justice to three research questions (unless you’re writing a book). It certainly cannot do justice to 15 research questions, which is often what those of us conducting VE assessments are asked to do.

If you find yourself having to answer multiple research questions in the same report, and you only have a few weeks to write it, then what inevitably happens is you find a few pieces of evidence—often anec-

dotal—to answer each question. In other words, the evidence for each question is thin and the knowledge gaps section becomes a laundry list of the need to develop further evidence to most of the research questions in your own assessment. There is not time to systematically and thoroughly answer every single question in a report that asks more than one or two questions—not to mention that if you **did** answer every single question in a rigorous manner, you would probably pen a 1,000-page treatise and no one would read it.

Keeping research questions simple and clear and the number of questions to no more than two is how researchers go about building a **body of knowledge**. Academics call this a research agenda: an outline of research questions that will guide the study of a given topic. One research project answers one (or at most two) question(s). Then academics move on to the next related question that logically flows from and builds upon previous research. It is the accumulation of knowledge over time through this simple and systematic method of developing a research agenda that moves forward the world’s knowledge on most topics of our time—from climate change to medical research.

USAID has a similar strategy of building up an existing body of knowledge through learning agendas.¹ However, because learning agendas tend to be project-specific and systems and processes are not built to aggregate knowledge **across** projects, intentions of deeper knowledge growth remain stymied by a lack of systematic processes to share knowledge from learning agendas with other practitioners and donors outside of USAID. Meta studies are valuable; however, they cannot replace the good old practice of limiting research questions to one or two per project and producing a coherent and systematic research agenda due to their post-hoc nature.

In sum, the first research principle is to generate questions that ask “who,” “what,” “where,” “why,” or “how” which facilitate describing, explaining, predicting, or proscribing. The second principle is to only pose one or two research questions per project and develop a plan for follow-on research that builds up a body of knowledge over time.

DEFINING TERMS

Properly understanding complex phenomena requires building up a body of knowledge. Take high blood pressure, for example. The causes of high blood pressure are numerous and complex; the effectiveness of individual or combined treatments for high blood pressure—behavioral changes, diets, and medications—are equally numerous and complex. Scientists and physicians have come to develop a sophisticated understanding of the causes of high blood pressure and what to do about it through hundreds of studies conducted over time that developed the evidence base needed to understand and address this health condition. Fortunately for those who suffer from high blood pressure, there is a common definition of what the phenomenon is and there is an accepted way to measure it. This facilitates the study of

1 Learning Lab, USAID, accessed September 1, 2019. <https://usaidealarninglab.org/>

causes and treatments by systematizing the definition and measurement of the disease across the vast research body.

Because VE and CVE are complex social phenomena, CVE researchers have struggled to find common definitions.² The term **terrorism** alone has over 250 definitions.³ This multiplicity of definitions makes accumulating a body of knowledge about the causes and treatments of the disease—in this case, VE—far more difficult for at least three reasons: 1) CVE researchers and practitioners hail from multiple disciplines and define terms based on their respective fields; 2) how Westerners define VE may not be relevant to rural villages with active violent extremist organizations (VEOs); and 3) governments and civil societies around the globe define VE and CVE differently. This section discusses each issue in detail, then offers a way forward to address it.

First, researchers tend to approach VE and CVE based on assumptions, definitions, and theories by experts who import concepts and research designs from their respective fields.⁴ The result is a variety of definitions and theories derived from different disciplines, rather than one interdisciplinary definition and theory. Similarly, CVE programs—such as those directed at increasing social cohesion, legitimacy, and resilience—are equally difficult to define in a uniform fashion across disciplines and projects.

Second, because most of the concepts, theories, and measurements are developed by experts (usually from Western perspectives), they may or may not be relevant to realities on the ground, thus they may not be valid. For example, experts may suggest that a reduction in VE should be measured in violent attacks or deaths, where fewer attacks would suggest a reduction in VE. However, when local communities in Afghanistan were asked what indicators they use to understand a reduction in VE, they offered responses such as “the ability of women and girls to move about safely in the community”⁵ and the presence of satellite dishes for television. Third, national governments across the globe define VE and CVE differently in accordance with their own legal frameworks and priorities; even national agencies within the same government adopt different definitions of these terms. Civil society actors also tend to define these concepts differently.

Fortunately, social scientists are not strangers to a lack of consensus about defining research concepts for complex social phenomena. Social scientists have long struggled to commonly define concepts such as democracy, war, insurgency, legitimacy, trust, etc. They deal with definitional ambiguity through paying careful attention to situating their definitions within a line of established scientific inquiry that similarly defines the concept. Good social science always clearly defines the key concepts of research, a practice that CVE research should adopt and make into a best practice. Conversely, bad social science does not

2 For further discussion of definitional issues, see J.M. Berger’s contribution to this edited volume: J.M. Berger, “Researching Violent Extremism: The State of Play,” RESOLVE, June 23, 2019, <https://resolvenet.org/research/researching-violent-extremism-state-play>

3 See Alex P. Schmid, ed., *The Routledge Handbook of Terrorism Research*. London, UK: Routledge, 2011, Appendix 2.1 “250-plus Definitions of Terrorism.”

4 Belquis Ahmadi and Eliza Urwin, “Measuring Peace and Violent Extremism,” *U.S. Institute of Peace*, March 16, 2018, 1.

5 Ahmadi and Urwin, “Measuring Peace and Violent Extremism,” 2.

clearly define terms, and often suffers from conceptual slippage – when the study begins by referring to a concept in one way, but then shifts the concept multiple times throughout to the point that, at the end of the study, the concept is completely different.

Applied physical scientists who study topics that obey natural laws and laws of physics have the National Institute of Standards and Technology to guide their practice in a systematic way.⁶ While we are unlikely to ever enjoy this level of systemization of our definitions due to the complex social nature of our research foci, the State Department offers some level of standardization of US-based foreign assistance programs (including peace, security, democracy, and governance, where CVE programs typically get categorized).⁷ However, this does not account for how other donors define their foreign assistance, or for how other government agencies define these terms. As a practice, we should strive to come closer to standardization because how we define our terms largely determines how we measure them and, therefore, how the findings add up to a body of knowledge on the phenomenon at hand.

RESEARCH DESIGN

Once the research question(s) and definition(s) are clear, the research design should follow. Research design refers to the researcher’s plan for each element that will be used to examine the research question, including definitions, measurement, data collection, data management, analysis, and reporting. It is essentially the recipe for how the researcher plans to go about conducting the study from start to finish. These elements of research design are nearly universally articulated in published quantitative research. However, in qualitative research, it is hard to find the same level of detail on how the researcher went about collecting, measuring, and analyzing data. As Zeev Maoz wrote over 15 years ago about qualitative case studies:

There is nearly a complete lack of documentation of the approach to data collection, data management, and data analysis and inference in case study research. In contrast to other research strategies in political research where authors devote considerable time and effort to document the technical aspects of their research, one often gets the impression that the use of case study absolves the author from any kind of methodological considerations. Case studies have become in many cases a synonym for free-form research where everything goes and the author does not feel compelled to spell out how he or she intends to do the research, why a specific case or set of cases has been selected, which data are used and which are omitted, how data are processed and analyzed, and how inferences were derived from the story presented. Yet, at the end of the story, we often find sweeping generalizations and ‘lessons’ derived from the case.⁸

6 National Institute of Standards and Technology, U.S. Department of Commerce, accessed September 1, 2019, www.nist.gov

7 See the Standardized Program Structure and Definitions website at <https://2009-2017.state.gov/f/c24132.htm>.

8 Zeev Maoz, “Case Study Methodology in International Studies: From storytelling to hypothesis testing.” In *Evaluating Methodology in International Studies: Millennial reflections on International Studies*, ed. Frank P. Harvey and Michael Brecher (Ann Arbor: University of Michigan Press, 2002), 455-75.

The absence of these details in research makes it impossible to evaluate the merits of the research design and, therefore, the validity of the findings. And so, we are left with a lot of dubious research in our practice that probably should not contribute to a body of knowledge on VE. The following sub-sections on data collection, measurement, and inference offer ways to guide researchers to both improve research and to more systematically write about their research and findings.

Data Collection

What kind of data do you need to collect to answer your research question and how are you going to collect it? That is really the essence of the data collection plan. Here are a few issues that have plagued field researchers to keep in mind.

Pitfall #1: Failure of the data collection tools. Data collection tools, such as surveys are no longer just the traditional pen and paper methods of olden days. Technological advances have made deploying surveys electronically much easier, which can vastly boost response rates (especially if the data is needed within a short timeframe). Surveys can be emailed, filled out on a website, collected over the phone, and deployed on mobile devices. However, researchers need to make sure that if they are going to use technology to deploy the survey, the technology works. Do mobile devices have sufficient battery power? How will they be charged? Does the internet work? With regards to the population, are the participants literate? Is the survey written in a language that the participants speak, read, write, or otherwise understand? Is the language spoken amongst the population a **written** language? If not, how will you deal with administering the survey? If the data collection tool is an interview or focus group discussion (FGD), do the data collectors have the materials needed to capture this data, such as pens and paper and functioning Dictaphones to record the content?

Pitfall #2: Enumerators misbehaving. Enumerators,⁹ sometimes intentionally but more often unintentionally, make subtle choices in the field that substantially affects the respondents' answers, including altering the wording of the questions, probing for specific responses, or asking guiding questions. In one field research project I conducted in Senegal, the research assistant (RA) I hired to conduct the FGDs argued with the participants and demanded that they change their opinions. In these situations, oversight of the enumerators or research assistants is critical. Because I observed the first FGD, I was able to have a polite conversation with the RA to let him know how I needed him to conduct the remaining discussions and we only lost the data from the first group. I also learned that I needed to train enumerators that I had not previously worked with, no matter their self-reported previous enumeration experience.

9 *Enumerators* are the people hired to go out and collect data in the field if the researcher does not collect the data themselves.

Pitfall #3: Inability to keep track of respondents. If the data collection tool is a panel survey or interview, i.e., a survey or interview that asks the same respondents questions over time, then researchers have to keep track of the respondents, which poses a pervasive problem. It is essential to have a solid plan for keeping in touch with respondents prior to deploying to the field, including keeping respondents' contact information in a safe and secure place.¹⁰ Researchers need to consider how they will guard the privacy and anonymity of the respondents, especially for respondents participating in sensitive research topics such as VE. The best practice here is to keep a document that codes the respondent to a number, e.g., "Jane Doe = respondent 1." This document should be kept separate from any data collected on the respondent. In the data itself, only the label "respondent 1" would be associated with the responses. If the response is reported, it should be cited as "Respondent 1." See the ethics section below for more on this, but if in doubt, consult the local Institutional Review Board (IRB) for the best practices in guarding respondents' identities.

Finally, because research on VE and CVE can be extremely sensitive, respondents may give false identities to the researchers. This can pose a practical challenge to those who need to provide evidence to their funders that they contacted real people for the study. The best way I have found to deal with this is to ensure that respondents understand my own process for guarding their data (communicated through informed consent – see more below in the ethics section) and to provide them with a written document, called a "Participant's Bill of Rights." This "Bill of Rights" clearly outlines what their rights are as participants in the study and provides contact details for superiors and the delegating institution so that they can report any violations of their rights that occurred in the study.

Pitfall #4: Failure to consider potential bias in the data collection. The way data are collected can itself cause bias. Who asks the questions, to whom we ask questions, how questions are asked, response options, and question order are common ways the act of collecting the data can skew findings or undermine data collection efforts.

First, who asks the questions is important. I'm a westerner working in West Africa. When I ask local community members questions, they may be less forthcoming to me than they would if someone from their own country asked them the same question. Therefore, if I want to collect data from local community members, I will opt to hire local enumerators to gather the data rather than doing so myself. However, if I want to collect data from West African officials, I might opt to do it myself because officials may be more forthcoming with me than they would with one of their constituents (or they may hope to improve relations with the US government through participation in the study, etc.). Researchers need to think through what potential biases may be created based on **who** is collecting the data and then be able to either mitigate it or caveat the findings. Researchers should always document bias.

¹⁰ Adapted from Dean Karlan and Jacob Appel, *Failing in the Field: What we can learn when field research goes wrong*, (Princeton, NJ: Princeton University Press, 2016)

Second, whom will you ask questions? How will you recruit participants and verify their identity to validate their inclusion? When and where will you meet with your study participants to collect their data? Put another way, what is the process for study inclusion? Here are some considerations you will want to make as you create the study inclusion plan:

- If you intend to conduct a survey, will you use a representative sample or a snowball sample? Will you survey the general population or officials?
- If you intend to conduct interviews, will you interview officials, former combatants, security services, civil society, students, etc.?
- Where will you gather data? Will you remain in the capital and ask elites their opinions about why some communities in their country are vulnerable to VE, or will you include data collection with residents of those communities?
- Will you only collect data from vulnerable communities or will you include more resilient communities?
- What are the implications for analysis of each of these decisions?

In creating the plan for study inclusion, researchers need to think through what kind of data they can get from different types of respondents and then make sure that they have adequate representation across respondents when the data collection is complete. Researchers should also think through how decisions about study inclusion will impact the data analysis and findings.

Finally, **how** the researcher asks the question can impact the data collection. To help explain why this is an issue, let's take a concrete example. Say you want to know about a community's grievances. One way to go about this is to create a close-ended survey that asks, "In your opinion, what are the greatest challenges facing this community" and then offer a list of options, e.g., "access to education," "security," "access to electricity," "access to healthcare," etc. Giving options that participants can choose from has the benefit of easing the data analysis because the responses are already provided and easy to code.

However, giving options means that you, the researcher made *a priori* decisions about grievances existing in the community. What if you miss the actual grievances because you forced respondents to choose from your list of potential grievances? The way to get around this potential omission is to ask open-ended questions. Here, we ask the same question but don't offer a list of options. This allows respondents to articulate grievances in their own words. We might discover that, actually, the grievance is that local officials don't speak local languages, so the community can't communicate with their own local government (this is a common grievance we have heard in the Liptako-Gourma region of the Sahel). The drawback of this approach is that it makes data analysis more onerous because the open-ended responses—which are qualitative in nature—need to be coded, which is a time-consuming and labor-intensive process.

Question order can also impact the data collection. In a survey, if you first ask about grievances and then immediately ask about satisfaction or trust in government, you are likely priming the respondent to give more negative responses than they may actually have, because you've just asked them about their problems and then you asked them to rate their government. There are volumes written about the nuances and issues with question order in social science research; suffice it to say, every detail of the question ordering and wording needs to be carefully considered and decisions intentionally weighed, documented, and caveated.

The key takeaway for the basic principles of data collection is to have a well-laid plan with contingencies for when things inevitably go wrong in the field. If the data collection has hiccups in the field, it does not necessarily mean the research needs to be entirely scrapped. It means that in the report researchers need to detail any issues that arose with data collection, document any potential bias it caused in the data itself (see more on bias below), and caveat the findings accordingly.

Research Ethics

The Declaration of Helsinki (1964)¹¹ established international ethical standards for medical research and serves as the foundational guideline for conducting research that involves people. Originally, the Declaration pertained to medical research. Social science experiments such as the Milgram Experiments (1963)¹² and the Stanford University Prison Experiment (1973)¹³ raised considerable ethical violations taking place in the social sciences and in the United States. Such experiments led to the Belmont Report (1974),¹⁴ which set out national ethical guidelines for any research involving human subjects, which also extend to research on animals. Today, any research conducted using US federal money that is subject to 45 CFR part 46—known as the Common Rule—is legally required to be reviewed by an Institutional Review Board (IRB) for ethical compliance. Most countries around the world that are signatories to the Declaration of Helsinki have their own IRBs; researchers need to check the ethical guidelines and compliance laws within their own countries and the countries of study.

Most research involving people is legally bound to comply with the international and national regulations of research ethics. More importantly, practitioners of CVE should be particularly sensitive to respecting the rights of study participants and should weigh the risks and benefits of having people participate in research. CVE research is sensitive and can put lives at risk. Furthermore, many of the people in our research have already gone through unspeakable trauma, and we should not inflict more on them in the name of research. Therefore, CVE researchers should carefully consider the potential risks to which we subject our research participants. Readers can learn about ethical guidelines through their local IRB or

11 "The Nuremberg Code," Ethical Principles for Medical Research, accessed September 1, 2019, <https://sites.jamanetwork.com/research-ethics/index.html>

12 Saul McLeod, "The Milgram Shock Experiment," *Simply Psychology*, 2017, <https://www.simplypsychology.org/milgram.html>

13 Saul McLeod, "The Stanford Prison Experiment," *Simply Psychology*, 2018, <https://www.simplypsychology.org/zimbardo.html>

14 "Belmont Report," U.S. Department of Health and Human Services, accessed September 1, 2019, <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>

through exploring the US Department of Health and Human Services online resources.¹⁵ Here, I outline a few of the more common ethical issues that arise in CVE research.

First, researchers need to consider ethical issues around study recruitment. Does the study in any way compel participation? For example, if you are interviewing former combatants in prison, are they compelled by the prison to participate? If so, that is a potential violation of their rights. What is the compensation for participating in the study? Is the compensation so high that it becomes a source of compulsion to join the study? Second, what are the risks and benefits to participants for engaging in the study? Do the benefits outweigh the potential risks to your participants? Third, depending on the study location, will the participants be overheard by others and does this pose ethical (or validity) issues? If you are working in communities experiencing an active VEO presence, how will you ensure VEOs do not find out who your study participants are? Field research conducted inside communities under active VE threat are vulnerable, and to keep both the research participants and enumerators safe, these questions need to be seriously considered. Fourth, as discussed above, how will the researcher safeguard the identity of participants?

Obtaining informed consent from each study participant is the best way to ensure your study remains ethically above board. Informed consent is a document that you and the study participant each sign. This document clearly states your information, your institution, your qualifications, the funder of the study, how the data will be processed, stored, analyzed, and reported, what the risks and benefits are to the participant, what the participants rights are (e.g. they have the right to not participate in the study or leave the study at any time without penalty, etc.), and provides contact information to your superiors and your institution so that the participant can report any ethics violations or negative consequences of participating in the study. You should also always obtain permission to use photographs of people. If you take photos of children, you should obtain their permission and their parents' or guardians' permission.

As with the other research basics mentioned in this chapter, it is incumbent upon us to think through these ethical considerations and have a well-laid plan to address them. Failure to adhere to ethical standards can have repercussions that reverberate not just for the study that violates ethics, but for future studies and programs and individual lives. Be assured that any ethical violations will contaminate the research pool and future researchers will not have the same access to data. It will also sully any programming that goes into the community based on the research findings. In the worst cases, it may adversely impact your study participants or the larger community.

Measurement

Some readers must be thinking: “but I do qualitative research. I don’t measure things.” I respectfully beg to differ. As King, Keohane, and Verba wrote over 25 years ago, qualitative and quantitative approaches

15 “IRB Registration,” U.S. Department of Health and Human Services, accessed September 1, 2019, <https://www.hhs.gov/ohrp/register-irbs-and-obtain-fwvas/irb-registration/index.html>

are simply “two styles of research, one logic of inference.”¹⁶ Setting aside the advances in qualitative research where concepts are now measured through coding or set theory analysis,¹⁷ the basic principles of measurement apply to qualitative research. This is because the basic principles of measurement rest on a foundation of logic. Logic applies to all research, quantitative and qualitative alike.

Ideally, CVE practitioners would have a unified set of metrics to apply across similar types of programs to build a body of knowledge that uses definitions and measures key concepts the same way. However, the field is not there yet.¹⁸ In the absence of agreed definitions for the core concepts of VE and CVE, exercising care about how we measure the definitions we use is especially important. The main principles of measurement are validity and reliability that, when adhered to, help standardize research practice. Failure to think through measurement issues often results in researchers falling back on measuring outputs, such as number of beneficiaries, or proxies of outcomes, such as community cohesion, support for the use of violence, or anti-American sentiment. Such proxies typically leave CVE practitioners little insight into a beneficiary community’s actual experience with VE.¹⁹ Furthermore, without a unified set of metrics, it is virtually impossible to provide evidence to support theories of what does and does not work under different contexts.

VALIDITY AND ACCURACY

Validity, or accuracy, describes the closeness of a measurement to the true value. This is the basic idea behind construct validity, which I discuss below. Two other important types of validity for our research practices are internal and external validity. This section discusses all three types of validity and issues we face in measuring VE and CVE.

CONSTRUCT VALIDITY

Construct validity means that the research tool measures the concept you want it to measure. It lies at the heart of one of the biggest challenges to the CVE field: measuring what VE is and what it means to counter VE. Our research and evaluation questions are almost always about a behavior, i.e. committing acts of terrorism, violent events, or the process of becoming radicalized. Yet, most of what our field considers cutting edge research does not measure behavior. Instead, we rely on attitudinal measurements—often about what respondents think they will believe in the future—in surveys and list experiments, such as whether respondents support a VEO or think that violence is sometimes justifiable; sometimes we simply rely on anti-American sentiment. While this may serve as a proxy measure (we need to do more research to evaluate this) in the absence of measuring actual behavior, we need to always caveat

16 Gary King, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific inference in qualitative research*, (Princeton, NJ: Princeton University Press, 1994), 3.

17 See, for example, King, Keohane, and Verba, *Designing Social Inquiry*; and Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago, IL: Chicago University Press.

18 Peter Romaniuk, “Does CVE Work? Lessons Learned from the Global Effort to Counter Violent Extremism,” *Global Center on Cooperative Security*, September 2015, 35–36.

19 Ahmadi and Urwin, “Measuring Peace and Violent Extremism,” 1–2.

such research and be straightforward about the extent of its meaningful contribution to explaining VE behavior.

Measuring attitudes, however, is wholly appropriate for attitudinal questions in our field. For example, we may want to know: What are people’s beliefs in X? What are citizens’ perspectives of the government and other power brokers—including VEOs—in their country? What are citizens’ grievances? Yet, measuring attitudes about these topics is difficult. First, people may interpret attitudinal questions differently. Even with surveys, perhaps the most standardized and systematized tool for data collection in CVE research, it is difficult to ensure that respondents will interpret even the simplest questions the same way. This is even more problematic when we add in language translations of survey responses. Study inclusion, who asks the questions, and how the questions are asked (see discussion above) can also affect responses to attitudinal questions and therefore the measurement.

Second, attitudes change. As Zaller and Feldman point out, survey research assumes that respondents have reasonably well-developed opinions for the questions posed to them and that “the standard view (of public opinion research) is that when survey respondents say they favor X they are simply describing a preexisting state of feeling favorably toward X.”²⁰ However, changes in responses to the same question when asked repeatedly, responses that change when trivial changes are made to question or responses, reactions to the context within which the questions are posed, and changes in responses when the question ordering is changed call this critical assumption into question. These behaviors suggest that respondents provide answers off the top of their heads. Asking a respondent the same question days, weeks, or months apart—while she can be influenced by events in her environments—can illicit inconsistent responses, making the measurement unreliable and invalid.

Third, surveys may not be internally invalid (see definition below) in the event that the questions are idealistic—based largely on Western democratic ideals—that do not capture valid attitudes. Another challenge to the internal validity of surveys (as well as interviews and FGDs) is human psychology. For example, people tend to want to give the answers they think will please others (social desirability bias).

These issues raise the need for survey creators and enumerators to have formal training in the psychology of surveying to know how to deal with them. Experience with surveying is not a substitute for this critical training; instead, it more likely guarantees that researchers will rely on shortcuts and employ poor instruments.

INTERNAL VALIDITY

Internal validity refers to whether the results from research are valid (true) for the study subjects. While we normally think about internal validity in experiments or quantitative research, it is equally important to consider in qualitative research. Qualitatively, we often rely on interviews and FGDs with officials

20 Stanley Feldman and John Zaller, “The Political Culture of Ambivalence: Ideological Responses to the Welfare State,” *American Journal of Political Science*, 36(1) 1992: 268-307.

and civil society organizations (CSOs). For example, when conducting VE assessments to inform program design, we tend to contact informants that our donors or governments know. Having conducted dozens—if not hundreds—of these interviews and FGDs, I am skeptical that they are internally valid. Study subjects are intended to be populations vulnerable to VE or VEOs, but, this technique does not actually collect information from vulnerable populations. Instead, the method collects information from elites about their opinions of vulnerable populations.

We hear over and over how it is youth and poor people who are vulnerable, with very little evidence or firsthand knowledge to support this assertion. Consequently, donors create programs for youth and poor people based the opinions of elites or accessible populations on their vulnerability, when, in reality, youth and poorer populations may or may not be vulnerable. Moreover, individuals may not even be the right unit of analysis; we may need to be looking at vulnerable communities instead. Overreliance on elites for data collection can render the research internally invalid—i.e., the results from the research do not hold for the study subjects (vulnerable populations) we want to know about. Put another way, we simply have the wrong study subjects in our research.

For example, I was once on a positive youth development program working with poor, uneducated youth in Tunisia. These youth had been identified as a population vulnerable to VE through the assessment that preceded the program design. That assessment relied on desk study of existing studies (which almost entirely consisted of elite interviews and FGDs with CSOs) and a few additional elite interviews, mostly with other implementers. After we started the program, the Tunisian Center for Research and Studies on Terrorism (CTRET) published a study that reviewed the court documents from a sample of one thousand alleged terrorist in the Tunisian court system to determine the demographic characteristics of these alleged terrorists.

Contrary to what the interviewees in our assessment told us—that the vulnerable population in Tunisia were poor, uneducated youth—this study showed that 40 percent of the alleged terrorists held university degrees and an additional 33 percent held secondary diplomas.²¹ Although the age category of youth is often so broad to include anyone aged 15–35 or 15–40, the study revealed variation across smaller age categories: 204 people were between the ages of 18 and 24, 275 were between the ages of 25 and 29, and 243 were between the ages of 30 and 34. Most of the thousand people in the study were men (965), although there were also 35 women. More than half (536) were single, 24 percent (239) were married, and 1 percent were divorced or widowed. What this meant for the Tunisia project was that the focus on the poorest, least educated communities was missing a large part of the vulnerable population, which included many middle-class and educated Tunisians.

The difficulty we face as practitioners and researchers is that we conduct assessments to identify the vulnerable populations for our programming. If we do not know who the vulnerable population is, how can

21 “A large proportion of terrorists in Tunisia are university graduates (study),” Agence Tunis Afrique Presse, October 26, 2016, <https://www.tap.info.tn/en/Portal-Society/8362201-a-large-proportion-of-terrorists-in>

we include them in our assessment? In this way, we face a sort of chicken-and-egg problem: we want to include data from vulnerable populations to design adequate programming, but we need to conduct the assessment to first identify the vulnerable population. Despite this challenge, what the previous Tunisian example suggests is that we can think through more critically who are our study subjects are and how to gather different kinds of data to ensure internal validity.

EXTERNAL VALIDITY

External validity is also known as generalizability, meaning the results from the sample hold true for the broader population. Though this principle is usually assigned to the realm of quantitative research, it is relevant for qualitative research as well. For example, if we conduct an FGD with youth in a certain neighborhood, we need to be careful to either evaluate whether the findings are valid for the whole neighborhood or we need to attribute the findings just those who participated in that specific FGD.

External validity is a difficult principle for CVE practitioners because so much of VE—and therefore research on VE—is hyper-localized. What researchers may find driving VE in one village may not be the same thing driving VE in a neighboring village. The implication is that the response to VE must be tailored to address the specific drivers across different subnational units. Therefore, many of our research findings are not generalizable. This is not a critical failure for our research. Rather, it points to the need for researchers to be very precise about how to communicate findings and to avoid making sweeping generalizations.

RELIABILITY AND PRECISION

Precision refers to the closeness of agreement among a set of measurements. Researchers often use it synonymously and interchangeably with reliability. Both reliability and precision mean that if the measurement is taken again—say, by another researcher—the values will be the same. In the attitude discussion above, reliability issues arise when the measurements of attitudes change over time. In qualitative research reliability and precision mean that if an interview is conducted again or conducted by a different enumerator, the evidence will be the same. Reliability can even be extended to the types of informants. That is, if the same kinds of informants are interviewed, will their information be about the same? Following, documenting, and sharing the data collection, measurement, and analysis plan is the best practice for achieving reliability.

Inference

What logical conclusions do we draw from the evidence? This is the basic principle of inference. In research on VE and CVE, there are three common kinds of inference flaws: relying on outputs to infer about outcomes; using attitudinal measurements to infer about behavior; and using evidence from one level of analysis—such as country context—to infer about another level of analysis—such as individuals.

In monitoring and evaluating (M&E) CVE programs, the tendency to rely on outputs (such as number of participants in a training) rather than on the program interventions' outcomes (such as the effect the

intervention had on reducing VE) results in lacking a strong evidence base about CVE impact. The reliance on outputs in USAID is partly a function of responding to the donor because CVE programs inevitably include a few foreign assistance indicators (aka F-indicators)²² that are almost solely output indicators. This is largely because measuring outputs is the only way to standardize measurement of foreign assistance. However, outputs do not show impact for complex social phenomenon such as VE and responses to VE. Therefore, we cannot say that because we trained a certain number of beneficiaries, we had an impact on VE. In our practice, we have to move far beyond reporting the number of beneficiaries or services and toward measuring outcomes.

As discussed above, our reliance on attitudinal measurements—instead of behavioral ones—in research on CVE means there is little evidence about what changes behavior. Put another way, just because we may evoke an attitudinal change does not mean we had any impact on behavior. We cannot use evidence about attitudes to infer about behavior. At the end of the day, it is behavior that needs to change to effectively counter VE. As several scholars have pointed out, people can hold radical attitudes without ever engaging in violent or radical behavior and not all terrorists radicalize.²³

Finally, ecological fallacy, or making inferences from data or evidence at one unit of analysis (usually a context measurement) to draw conclusions about another unit of analysis (usually individuals) is a common pitfall in our practice. In the early 2000s, a popular hypothesis was that poverty drives terrorism. This hypothesis was borne out of country level studies examining the correlation between national incidents of violence and a variety of national poverty measurements (such as GDP and inequality). Many national-level studies since then have demonstrated this is not the case²⁴. Nevertheless, inferences that poverty drives terrorism were extended to the individual level: poor people are more likely to be terrorists.

However, studies at the individual level have demonstrated that many people who commit terrorist acts come from middle class or prominent families and that they commit terrorist acts for a number of different reasons. “One of the most significant understandings gained from academic research over recent years is that individuals involved in terrorist activities exhibit a diversity of social backgrounds, undergo rather different processes of violent radicalisation and are influenced by various combinations of motivations.”²⁵ In sum, we cannot infer findings about one level of analysis based upon evidence from another level of analysis.

22 See Standard Foreign Assistance Indicators, U.S. Department of State, Archived, accessed September 1, 2019, <https://2009-2017.state.gov/f/indicators/>

23 Randy Borum, “Radicalization into Violent Extremism II: A Review of Conceptual Models and Empirical Research.” *Journal of Strategic Security*, 4(4) 2011: 37-62; Clark McCauley and Sophia Moskalenko, “Mechanisms of Political Radicalization: Pathways toward Terrorism.” *Terrorism and Political Violence* 20, 2008: 415–33; Fernando Reinares, *Radicalisation Processes Leading to Acts of Terrorism: A Concise Report Prepared by the European Commission’s Expert Group on Violent Radicalisation*, European Commission, 2008.

24 See, for example, James A Piazza, “Rooted in Poverty?: Terrorism, Poor Economic Development, and Social Cleavages,” *Terrorism and Political Violence*, 18(1) 2007: 159-177.

25 Reinares, *Radicalisation Processes Leading to Acts of Terrorism*, 11; see also McCauley and Moskalenko, “Mechanisms of Political Radicalization”; Borum, “Radicalization into Violent Extremism II”; Schmid, *The Routledge Handbook of Terrorism Research*.

IMPLICATIONS FOR POLICY AND PRACTICE

To understand the implications of conducting poor research for policy and practice, it is instructive to borrow lessons learned from established fields. Take for example, the medical practice. For thousands of years, esteemed and charismatic people practiced medicine by prescribing treatments based on anecdotes, biases, and probably a lot of psychological traps, such as confirmation bias and cognitive dissonance.²⁶ Early advances in public health happened through policies aimed at combatting urban issues such as sanitation, housing, and infectious diseases—not by doctors. Only in the past 25 to 30 years, with the advent of better research technology—especially randomized experiments—has medicine moved to an evidence-based practice. Many old guard physicians rebuked this transition; only more recently, as those physicians left the field, has evidence-based practice really taken hold.²⁷

While evidence has moved the medical field forward, it also shed light on **what not to do**. According to Harvard health care policy professor Anupam Jena, about 15% of the time science leads to medical reversals.²⁸ The inconvenient fact is that the human body is a complex organism, which makes diagnosing and treating health problems difficult. Similarly, CVE deals with complex social phenomena within multiple complex systems, which makes it more challenging to investigate. We need to follow the standard practices of evidence-based research to generate the knowledge base needed to get a better handle on how to intervene effectively in response to VE and to know what not to do.

For CVE, experienced policymakers and senior military professionals have mismatched skillsets to these challenges. They cut their teeth on foreign policy during the Cold War era, when the great threat was another nation-state—a challenge that states have had practice dealing with for centuries. Trained to address state-level threats, senior policymakers have made underwhelming decisions on countering terrorism, a complex non-state actor problem. As the recent Congressional mandate to the United States Institute of Peace (USIP) to set up a task force on VE in fragile states demonstrates, the response to terrorism has not stopped VE from spreading around the globe.²⁹ The USIP Task Force findings and recommendations point to the need to do more prevention and less countering.³⁰ Time will tell whether high-ranking decision makers will head this call to change.

26 Confirmation bias is the tendency to interpret new evidence in a way that conforms to previously held beliefs. For example, if you believe your child is messy, each time they leave their clothes on the floor and toys strewn throughout the house will confirm that belief. Cognitive dissonance refers to the discomfort people feel when their beliefs contradict their behavior. For example, smokers may have cognitive dissonance around smoking. They may believe smoking is bad for their health and desire to quit but they continue smoking.

27 For an engaging discussion about this evolution, see the 3-part Freakonomics Podcast “Bad Medicine,” 2016, <http://freakonomics.com/podcast/bad-medicine-part-1-story-98-6/>.

28 See Freakonomics Podcast “Bad Medicine,” 2016, <http://freakonomics.com/podcast/bad-medicine-part-3-death-diagnosis/>

29 Task Force on Extremism in Fragile States, “Preventing Extremism in Fragile States: A new approach,” *U.S. Institute of Peace*, February 26, 2019. <https://www.usip.org/publications/2019/02/preventing-extremism-fragile-states-new-approach>

30 Task Force on Extremism in Fragile States, “Preventing Extremism in Fragile States: A new approach”

Meanwhile, across the board, government seems to increasingly respond better to evidence than experience or expert opinion. There is a growing drum beat from the U.S. Congress to show evidence-based impact of CVE programming. Most importantly, there is a real human need to craft better responses to help build peace for the people living under the siege of VEOs. Going back to research basics to provide the most solid and rigorous evidence possible, and systematically building a body of knowledge off the findings of this evidence base, is how researchers and analysts can do their part to improve the overall response to VE and contribute to a more peaceful world.

SOURCES

- Agence Tunis Afrique Presse. "A large proportion of terrorists in Tunisia are university graduates (study)." October 26, 2016. <https://www.tap.info.tn/en/Portal-Society/8362201-a-large-proportion-of-terrorists-in>
- Ahmadi, Belquis and Eliza Urwin "Measuring Peace and Violent Extremism," *U.S. Institute of Peace*, March 16, 2018.
- Berger, J.M. "Researching Violent Extremism: The State of Play," *RESOLVE*, June 23, 2019. <https://resolvenet.org/research/researching-violent-extremism-state-play>
- Borum, Randy. 2011. "Radicalization into Violent Extremism II: A Review of Conceptual Models and Empirical Research." *Journal of Strategic Security*, 4(4) 2011: 37-62.
- Ethical Principles for Medical Research. "The Nuremberg Code." Accessed September 1, 2019, <https://sites.jamanetwork.com/research-ethics/index.html>
- Feldman, Stanley and John Zaller. "The Political Culture of Ambivalence: Ideological Responses to the Welfare State," *American Journal of Political Science* 36(1) 1992: 268-307.
- Freakonomics Podcast. "Bad Medicine." 2016. <http://freakonomics.com/podcast/bad-medicine-part-1-story-98-6/>.
- Karlan, Dean and Jacob Appel. *Failing in the Field: What we can learn when field research goes wrong*. Princeton, NJ: Princeton University Press, 2016.
- King, Gary, Robert O. Keohane, and Sidney Verba. *Designing Social Inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press, 1994.
- Maoz, Zeev. "Case Study Methodology in International Studies: From storytelling to hypothesis testing." In *Evaluating Methodology in International Studies: Millennial reflections on International Studies*, edited by Frank P. Harvey and Michael Brecher 455-75. Ann Arbor: University of Michigan Press, 2002.
- McCauley, Clark, and Sophia Moskalkenko. "Mechanisms of Political Radicalization: Pathways toward Terrorism." *Terrorism and Political Violence* 20, 2008: 415-33.
- McLeod, Saul. "The Milgram Shock Experiment." *Simply Psychology*. 2017. <https://www.simplypsychology.org/milgram.html>
- McLeod, Saul. "The Stanford Prison Experiment." *Simply Psychology*. 2018. <https://www.simplypsychology.org/zimbardo.html>
- Mishler, William, and Richard Rose. "What Are the Origins of Political Trust?" *Comparative Political Studies* 34 (1) 2001:30-62.
- Piazza, James A. "Rooted in Poverty?: Terrorism, Poor Economic Development, and Social Cleavages," *Terrorism and Political Violence*, 18(1) 2007: 159-177.
- Ragin, Charles C. *Fuzzy-Set Social Science*. Chicago, IL: Chicago University Press, 2000.
- Reinares, Fernando. *Radicalisation Processes Leading to Acts of Terrorism: A Concise Report Prepared by the European Commission's Expert Group on Violent Radicalisation*. European Commission, 2008.
- Romaniuk, Peter. "Does CVE Work? Lessons Learned from the Global Effort to Counter Violent Extremism" *Global Center on Cooperative Security*, September 2015.
- Schedler, Andreas and Rodolfo Sarsfield. "Democrats with Adjectives: Linking Direct and Indirect Measures of Democratic Support." *European Journal of Political Research* 46(5) 2007: 637-659.
- Schmid, Alex P., ed. *The Routledge Handbook of Terrorism Research*. London, UK: Routledge, 2011.
- Task Force on Extremism in Fragile States. "Preventing Extremism in Fragile States: A new approach." February 26, 2019. <https://www.usip.org/publications/2019/02/preventing-extremism-fragile-states-new-approach>
- USAID. Learning Lab. Accessed September 1, 2019. <https://usaidealarninglab.org/>
- U.S. Department of Commerce. National Institute of Standards and Technology. Accessed September 1, 2019, www.nist.gov
- U.S. Department of Health and Human Services. "Belmont Report." Accessed September 1, 2019. <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>
- U.S. Department of Health and Human Services. "IRB Registration." Accessed September 1, 2019. <https://www.hhs.gov/ohrp/register-irbs-and-obtain-fwais/irb-registration/index.html>
- U.S. Department of State. Standard Foreign Assistance Indicators. Archived, accessed September 1, 2019. <https://2009-2017.state.gov/f/indicators/>

RESOLVE

NETWORK

INSIGHT INTO VIOLENT EXTREMISM AROUND THE WORLD

The RESOLVE Network is a global consortium of researchers and research organizations committed to delivering fresh insight into violent extremism around the world. The Network provides access to open-source data, tools, and curated research to ensure policy responses to violent extremism are evidence based. Members of the Network work in parts of Africa, Asia, Europe, and the Middle East to promote empirically driven, locally defined responses to conflict and to support grassroots research leadership on violent extremism.

Our partners operate in more than 25 countries where challenges with conflict are an everyday reality. We are passionate about amplifying credible local voices in the fight to mitigate the destabilizing risks of social polarization and political violence. The RESOLVE Network Secretariat is housed at the U.S. Institute of Peace, building upon the Institute's decades-long legacy of deep engagement in conflict-affected communities.

To learn more about the RESOLVE Network, our partners and how to get involved visit our website, www.resolvenet.org, and follow us on Twitter: @resolvenet.





**UNITED STATES
INSTITUTE OF PEACE**
Making Peace Possible

2301 Constitution Avenue, NW
Washington, DC 20037